Gene Expression, Vol. 4, pp. 63–75 © 1994
Printed in the USA.

1052-2166/94 $3.00 + .00

Bovine and Rodent Tamm-Horsfall Protein (THP) Genes: Cloning, Structural Analysis, and Promoter Identification

HERON YU,[1] FEROZ PAPA,[1] AND VIKAS P. SUKHATME[2]

Renal Division, Department of Medicine, Beth Israel Hospital and Harvard Medical School, 330 Brookline Avenue, Boston, MA 02215

We have isolated bovine and rodent cDNA and genomic clones encoding the kidney-specific Tamm-Horsfall protein (THP). In both species the gene contains 11 exons, the first of which is noncoding. Exon/intron junctions were analyzed and all were shown to follow the AG/GT rule. A kidney-specific DNase I hypersensitive site was mapped onto a rodent genomic fragment for which the sequence is highly conserved in three species (rat, cow, and human) over a stretch of 350 base pairs. Primer extension and RNase protection analysis identified a transcription start site at the 3′ end of this conserved region. A TATA box is located at 32 nucleotides upstream of the start site in the bovine gene and 34 nucleotides upstream in the rodent gene. An inverted CCAAT motif occurs at 65 and 66 nucleotides upstream of the start site in the bovine and rodent genes, respectively. Other highly conserved regions were noted in this 350 bp region and these are likely to be binding sites for transcription factors. A functional assay based on an in vitro transcription system confirmed that the conserved region is an RNA Pol II promoter. The in vitro system accurately initiated transcription from the in vivo start site and was highly sensitive to inhibition by α-amanitin at a concentration of 2.5 μg/ml. These studies set the stage for the further definition of *cis*-acting sequences and *trans*-factors regulating expression of the THP gene, a model for kidney-specific gene expression.

TAMM-HORSFALL protein (THP) is a glycoprotein with a molecular weight of 80 kDa, initially characterized by Tamm and Horsfall (1950) and subsequently purified from the urine of pregnant women by Muchmore and Decker (1985) as uromodulin. Pennica et al. (1987) showed that the two proteins were identical. THP is the most abundant protein in urine and is present in the kidneys of all placental mammals (Fletcher, 1972). Because one of our long-term research goals is to understand how gene expression is regulated in the kidney, we have chosen to use the THP gene as a model system for the following reasons. (1) The expression of THP is limited to the kidney in a nephron segment specific distribution. It is therefore a good example of a kidney-specific gene. Within the kidney, it is expressed only in the thick ascending limb of Henle's loop and in the distal convoluted tubule (Schenk et al., 1971). THP is produced in all mammalian species with a similar nephron-specific localization. The gene is therefore likely to have conserved *cis* regulatory sequences, and comparison of homologous regions from different species could implicate motifs that might bind tissue-specific transcription factors. (2) THP mRNA is present in high amounts (0.5–1% of message). The abundance of this message should greatly facilitate studies of gene regulatory sequences.

As a first step toward elucidation of the mechanisms that regulate kidney-specific gene expression, we have isolated and characterized the bovine and rodent THP genes. These animal homologues of the human THP gene are desirable due to the availability of kidney tissue that is required for chromatin structural studies and for making nuclear extracts

Received April 4, 1994; revision accepted June 22, 1994.
[1]Equal contribution.
[2]Address correspondence to Vikas P. Sukhatme.

(see below). In this article we describe the cloning of cDNAs containing the entire coding region and most of the noncoding region for bovine and rodent THP. The entire sequence of the cDNAs and the intron/exon structure of the two genes are also presented. We have identified a DNase I hypersensitive site in a highly conserved region that is also the promoter of this gene, as demonstrated by RNA mapping analysis and through an in vitro functional assay.

MATERIALS AND METHODS

Cloning, Restriction, and Sequence Analysis

cDNA libraries were in lambda gt10. The rodent library was a kind gift of Dr. G. Bell (University of Chicago). Standard procedures were used for cloning, restriction, and sequence analysis (Sambrook, 1989). Genomic libraries were in the lambda DASH II vector (Stratagene). cDNA and genomic inserts were cloned into pUC19 and pBluescript II SK, respectively.

RNA Isolation and Northern Analysis

RNA was isolated by the Chomczynski protocol (Chomczynski and Sacchi, 1987). Northern blotting and hybridization were as described (Current Protocols, 1991). Inserts from cDNA plasmids pBTHP1 and pRTHP1 were labeled by random priming (Feinberg and Vogelstein, 1983).

DNA Sequencing

The cDNA phage insert was subcloned in pUC19 and DNA sequencing was carried out using the chain-termination method with Sequenase (United States Biochemical).

DNase I Hypersensitive Site Analysis

Nuclei were isolated from 5 g of kidney and liver by mincing tissue in 25 ml of 0.25 M sucrose, 10 mM Tris HCl (pH 7.5), 5 mM NaCl, 3 mM $MgCl_2$, 0.5% NP-40. Tissue was homogenized in a Potter-Elvejem homogenizer and the nuclei were isolated by centrifugation. They were washed three times with the same buffer without NP-40 and resuspended at a DNA concentration of 1 mg/ml. DNase I was added to aliquots of the nuclei at 0, 1, 2, 5, and 10 μg/ml final concentration and incubation was carried out for 2 min at 25°C. One volume of stop solution (12.5 mM EDTA, 0.5% SDS, 500 μg/ml proteinase K) was added and incubation was carried out for 30 min at 37°C. Phenol-chloroform extractions were performed twice and a chloroform extraction was carried out before the DNA was precipitated with EtOH. The DNA was resuspended in TE and digested with Hind III, electrophoresed, and blotted onto nylon membranes. Prehybridization and hybridization was in 6 × SSC, 0.5% SDS, 100 μg/ml salmon sperm DNA, 5 × Denhardt's reagent (Current Protocols, 1991), and 50% formamide. A random printed, ^{32}P-labeled 300 bp intron 1 probe that was generated by PCR (diagrammed in Fig. 3) was used.

RNA Mapping Analysis

RNase protection assays were according to the method of M. Gilman (Gilman in Current Protocols, 1991). Total RNA (10 μg) was used for all reactions. For bovine THP, an EcoRI-Pst I 800 bp genomic fragment that contains the first exon was subcloned into pBluescript II SK+. The plasmid was digested with Ssp I and the 220 bp fragment was transcribed by T_3 RNA polymerase to give an antisense riboprobe. For rat THP, a PCR fragment from the rat conserved region −350 to +35 was cloned into the SmaI site of pBluescript II SK. The plasmid was cut with XhoI and transcribed with T_3 RNA polymerase to give an antisense riboprobe; 5 × 10^5 cpm were used for the hybridization.

Primer extension analysis was carried out as described by R. Kingston (Current Protocols, 1991) with specific exon 2 oligos BTHP4 and BTHP7 for bovine samples and RTHPRE36 and RTH5 for rat samples. mRNA was prepared by using oligo-dT columns (Boehringer Mannheim). Fewer nonspecific extensions resulted when 2 μg mRNA rather than 10 μg total RNA was used.

In Vivo Transcription

Extracts for in vitro transcription were made by the method of Gorski (1986) with the following modifications for kidney extracts: homogenization buffer 1 was supplemented with 1% nonfat milk, 5 μg/ml bestatin, 0.1 μg/ml leupeptin, 0.1 μg/ml pepstatin, 1 mM $NaMO_4$ (we found that inclusion of this phosphatase inhibitor was critical to the success of the procedure); homogenization buffer 2 contained the same supplement. Nuclear lysis buffer and nuclear dialysis buffer contained the same supplement but without the nonfat milk. Moreover, prior to homogenization, the kidneys were passed through a garlic press (Crate and Barrel worked best!).

The transcription reactions were essentially as

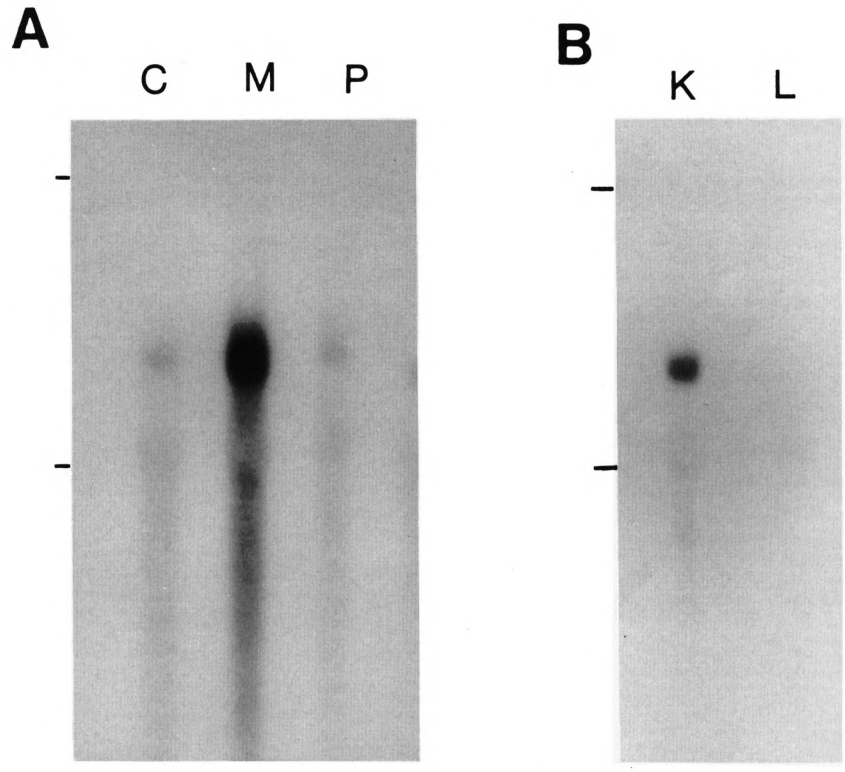


FIG. 1. (A) Northern blot of total cortical (C), medullary (M), and papillary (P) bovine RNA probed with the insert of pRTHP1. (B) Northern blot of total kidney (K) and liver (L) rat RNA probed with the same insert. Markers refer to 28S and 18S ribosomal sizes (≈5 kb and ≈2 kb, respectively).

described (Gorski et al., 1986). Template DNAs were prepared by a triple-spin $CsCl_2$ procedure and used at final concentrations of 50 μg/ml; 60 μg/ of protein was used per assay and α-amanitin was at 2.5 μg/ml.

RESULTS

Rodent and Bovine THP cDNA Homologues

A 550 bp BamH1-Hind III fragment from the human cDNA clone UM19 was used to screen rodent and bovine kidney cDNA libraries (Hession et al., 1987). Approximately 0.5% of plaques were positive in each library. The longest bovine clone, designated pBTHP1, had an insert of 2.5 kb whereas the longest rodent clone, pRTHP1, contained a 2.2 kb fragment. The inserts were used as probes for Northern analysis. As seen in Fig. 1A, the bovine probe hybridized strongly to a 2.6 kb transcript in outer medulla but weakly to cortex or papilla. The rodent probe hybridized to kidney but not liver (Fig. 1B).

pBTHP1 and pRTHP1 were fully sequenced (Fig. 2). pBTHP1 contains 1929 bp in its longest open reading frame and pRTHP1 contains 1932 bp leading to a predicted molecular weight of THP of ≈ 68 kDa (taking into account the 24 amino acidleader sequence cleaved from the precursor protein). Given the measured M.W. of ≈ 90 kDa, the carbohydrate moiety accounts for about 25% of the total. At the nucleotide level the bovine sequence shows 72.5% and 80.0% homology to the rodent and human sequences, respectively. Pennica et al. (1987) noted an unusually high number of CpG dinucleotides in the human third exon. This is also seen in the bovine third exon (63 CpGs). However, only 29 CpGs were noted in the rodent third exon, still a number greater than the predicted number due to CpG suppression, which is 18. The significance of this CpG cluster is not known.

Genomic Clones Spanning the Entire Rodent and Bovine THP Genes

Rat genomic clones were isolated by screening approximately 2×10^6 recombinant phage from a rodent genomic library (Stratagene) using the pRTHP1 cDNA insert as a probe. This screening yielded 16

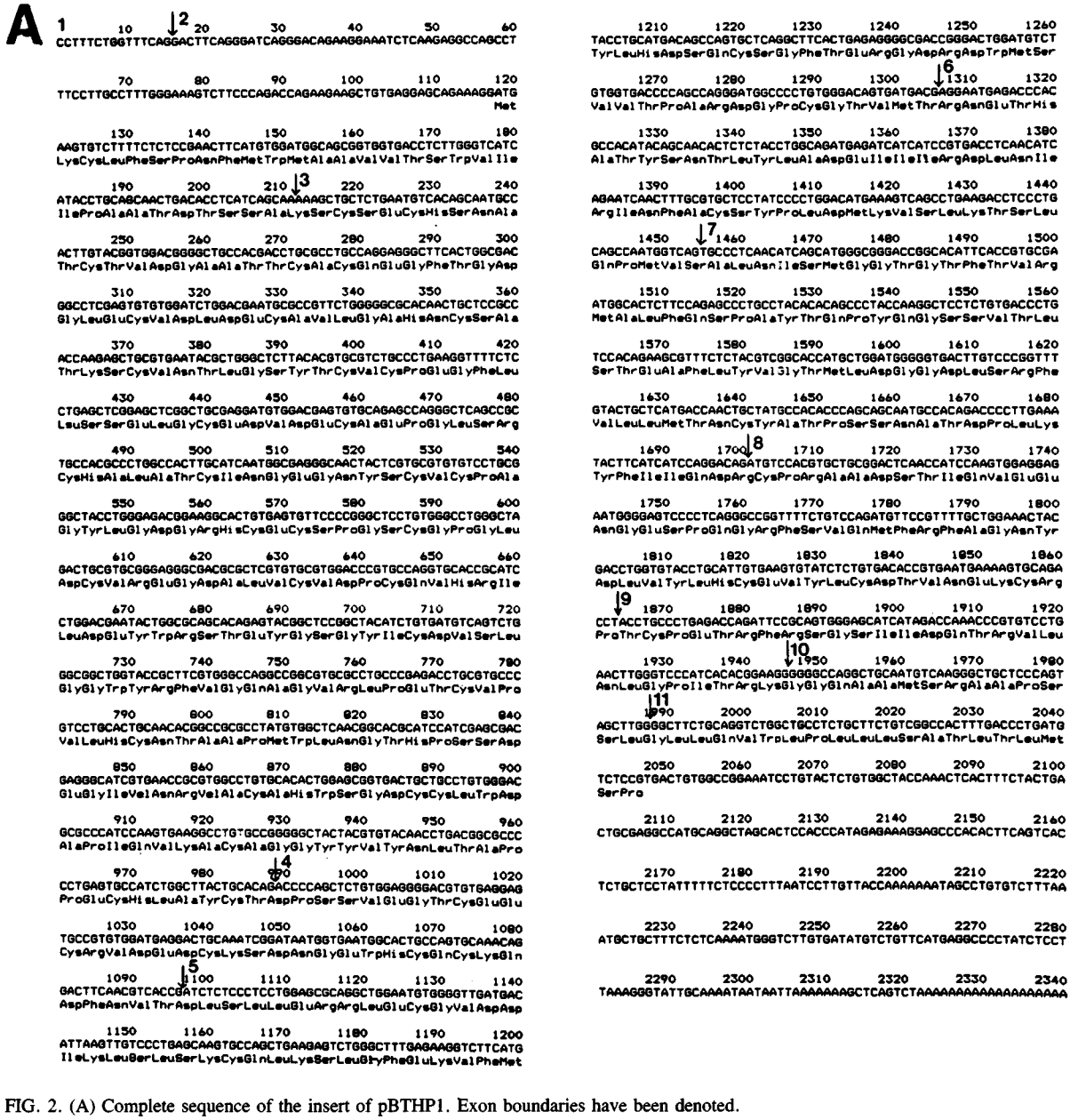


FIG. 2. (A) Complete sequence of the insert of pBTHP1. Exon boundaries have been denoted.

positive clones. The clones were further characterized and subgrouped by plaque hybridization with 17-mer oligos corresponding to selected exon sequences. Three clones were found to overlap and contained all 11 exons. Restriction blot analysis of these clones showed that the gene spans approximately 16 kb (Fig. 3B). The pBTHP cDNA insert was similarly used to screen 1×10^6 phage of a bovine genomic library (Stratagene). Out of eight positive clones, three were chosen for further structural analysis by restriction blot hybridization with four fragments of pBTHP1 (Fig. 3A).

The bovine gene spans 25 kb. Exon/intron junctions were assigned to both genes by sequence analysis using the junctions reported by Pennica et al. as a guide. All 10 exon/intron junction sequences follow the AG/GT rule (Breatnach and Chambon, 1981). The junctions show complete homology to the human sequences. Both the rat and bovine THP genes consist of 11 exons, the first of which is noncoding.

A Kidney-Specific DNase I Hypersensitive Site at the THP Gene Promoter

Because promoters are frequently associated with DNase I hypersensitive sites (Gross and Garrard,

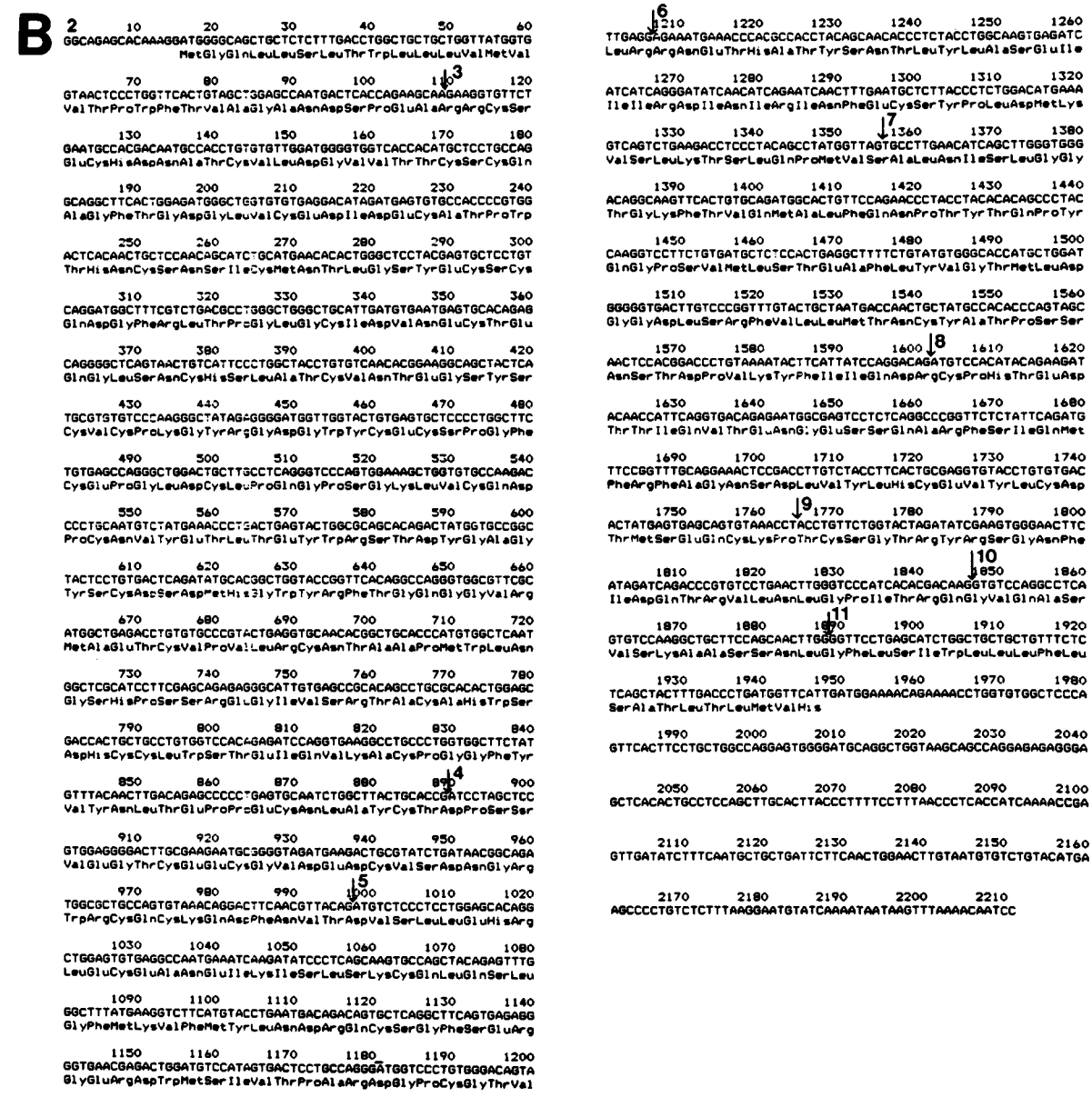


FIG. 2. (B) Complete sequence of the insert of pRTHP1.

1988), we searched for such sites in genomic fragments that were just upstream of the rodent cDNA 5′ end. Rat kidney nuclei were subjected to limited amounts of DNase I, then restricted for genomic Southern analysis. A genomic probe of 300 bp (derived by PCR) located in intron 2 was used (denoted in Fig. 3B as "P"). The expected 3 kb Hind III-Hind III fragment is reduced to 1.7 kb in a fraction of the nuclei (Fig. 4). This heterogeneity was anticipated because only about 10% of the cells in a whole kidney (those in the thick ascending limb and early distal nephron) are thought to express the THP gene. Importantly, the hypersensitive site is absent from rat liver nuclei subjected to the same manipulations and also from DNA isolated from a rat kidney directly (without an intermediate nuclear fractionation and DNase I digestion step). We sequenced the locus of the hypersensitive site and found striking homology between the rat sequence and sequence obtained from a region of a human genomic clone just upstream of the alternatively spliced exon "a" reported by Pennica et al. A rat probe spanning the hypersensitive site (derived through PCR) was then used to locate the homologous region in the bovine genomic clones. This sequence comparison is presented in Fig. 5. The conserved

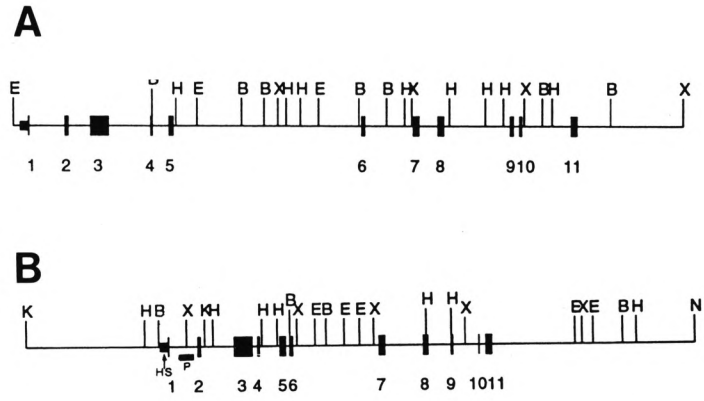


FIG. 3. Genomic organization of the bovine (A) and rat (B) THP gene. Boxes represent exons and have been numbered. The locus of the hypersensitive site in the rat clone has been marked as HS. Restriction sites are abbreviated: K, Kpnl; E, EcoRI; H, HindIII; X, Xbal; B, BamHI; N, Ncol.

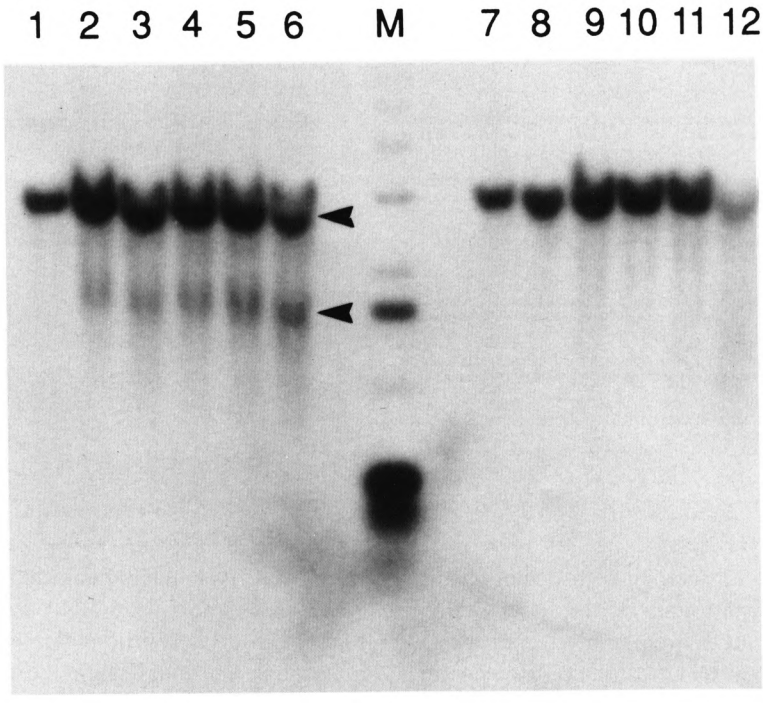


FIG. 4. DNase I hypersensitive site analysis. Lanes 1 and 7: DNA from rat kidney and liver (respectively) isolated without an intermediate nuclear purification step so as to minimize degradation. Lanes 2–6: DNA from kidney nuclei subjected to DNase I at concentrations of 0, 1, 2, 5, and 10 μg/ml. Lanes 8–12: DNA from liver nuclei subjected to 0, 1, 2, 5, and 10 μg/ml of DNase I. All DNAs were digested with HindIII. The probe is from intron 1, a nonrepetitive, PCR-generated 300 bp fragment (see the Materials and Methods section). The original 3 kb HindIII-HindIII band (top arrow) is reduced to 1.7 kb (lower arrow) in approximately 10% of the sample and only in samples derived from kidney nuclei. Lane M is a marker lane, with a "1 kb ladder" containing 1 kb marker DNAs and a Hinf I digest of pBR322 (Gibco, BRL).

region spanned approximately 350 bp in all three species and contained distinctive promoter hallmarks: a TATA box and an inverted CCAAT box (underlined).

RNA Mapping to Locate the Transcription Start Sites

Primer extension analysis was performed on bovine outer medullary RNA. Two oligo probes, BTHP4 and BTHP7, randomly chosen from exon 2 produced primer-extension products at 16 nucleotides upstream from nucleotide position 1 of the cDNA containing plasmid pBTHP1 (Fig. 6A). An RNase protection assay using a 220 bp fragment (Ssp I-Pst I), which contained the first exon, gave a protected nucleotide size at a range of 32–33 confirming the first exon size of 31 nucleotides as shown by the primer extension assay (Fig. 7A). When primer extension analysis was performed on rat kidney RNA using two different antisense primers RTHPRE36 and RTH5, separated by 57 bases, major extension products of 102 and 159 nt were detected (Fig. 6B). RNase protection analysis using a 385 bp antisense riboprobe spanning the conserved region (from nt 229 to 614, Fig. 5) gave a protected fragment of 30 bp (Fig. 7B). Liver RNA was negative by both methods. The results of the mapping studies in both species were in close agreement and identified a cap site 32 bp and 34 bp downstream of the TATA boxes underlined in Fig. 5.

In Vitro Transcription

To date, and despite extensive searching, we have not found a cell line expressing THP message. To begin to study cell type specific regulation of THP, we elected to attempt in vitro transcription from a kidney extract, initially derived from whole rat kidneys. Similar extracts from other organs (e.g., liver) have proven to be valuable reagents for studying determinants of tissue-specific transcription in vitro. Our kidney extracts retain the capacity to utilize the in vivo cap site and are sensitive to low concentrations of α-amanitin (2.5 μg/ml). We found that the production of transcriptionally competent nuclear extracts from kidney required protease inhibitors in addition to those described by Schliber (Gorski et al., 1988), and the phosphatase inhibitor sodium molybdate (see the Materials and Methods section).

A rat THP promoter fragment from −1600 bp to −3 bp (relative to the assigned start site in the rodent gene, Fig. 5) was fused just upstream of the G-less cassette in pC_2AT (a kind gift of Dr. R. Roeder) to give pTHP380. Rat kidney nuclear extracts faithfully initiated transcription from the in vivo cap site (3 bp into the cassette) to produce the expected 380 bp transcript that is truncated by the chain-terminating guanosine nucleotide analog 3′-*O*-methyl GTP (lane 1, Fig. 8) In addition, the inclusion of α-amanitin at 2.5 μg/ml resulted in complete poisoning of transcription as is seen with other RNA Pol II promoters (lane 2, Fig. 8). A construct that has the orientation of the promoter fragment of pTHP380 reversed also gave no signal (data not shown). The results of this functional assay provide further proof that the 350 bp conserved region is the promoter for the THP gene.

DISCUSSION

One approach to the study of renal development is to define a hierarchical set of renal transcription factors that participate in a temporal and spatial cascade of transcriptional activation of kidney-specific genes, ultimately leading to the differentiated state. A tissue-specific marker gene can serve as a necessary springboard in such an approach. For example, much insight into the problem of hepatogenesis has been gained by studying the mechanisms that underlie liver-specific activation of the albumin gene (de Simone and Cortese, 1988).

As a first step towards elucidation of the mechanisms of transcriptional regulation of a kidney-specific gene, we have cloned and characterized the bovine and rat THP genes. Bovine and rat cDNA-clones were obtained by screening kidney cDNA libraries with a human THP cDNA fragment. Genomic clones were then isolated by using the cDNAs as probes. The location of exons on the genomic map along with a detailed analysis of the exon/intron junctions was a necessary first step in delineating the boundaries of the THP transcription unit. The promoter was identified through a combination of strategies that included DNase I hypersensitive site analysis, interspecies sequence comparison, and RNA mapping studies. That the promoter fragment identified by these means directs RNA Pol II-mediated transcription in vitro further authenticates its central role in the transcription unit of the THP gene.

A cell line that expresses the THP mRNA at the normally high in vivo levels does not exist. Therefore, we have chosen to work with the bovine and rat genes to characterize gene regulatory elements by species comparisons, using transgenic mice and by biochemical approaches. The latter strategy requires large quantities of kidney nuclei that are readily available from animals. The cow kidney offers the additional advantage of further dissection of the

```
             *        *   **      *    **  *    *       *
bTHP AATTTCTTGATTCACAGAGCATCTGGTCCAATGATG TCTGAATTGCCTTCTGTCTCTGA   59
rTHP CTAGTCTTGTCTGACAGAG GTCCAGTTGAGGGATG TCCAGATGGTCT  TGCAACCGA   56
hTHP C AG    AGTGGGTC  AG GTCCAG TGATGTCTGAACTACCTTCTGG  TTCTGACTT   50

        * *  * *******   ** **    *    ** ***  * *  *  *
bTHP CCTTCAGGCATTCTCAGCTCCTTTCCTGCTCACATCGGGACCCCAGGGAAGCTGGT T     116
rTHP    TAA  CTTTCTCAGAGACTCTCTCTTTCCTGTCTGGACTCTAGTGGGGAGGAC T     108
hTHP    TCAGCCATTCTCAGCTCCTCTCTTGCTTGTGTCTGGATTCTAAGGCTGATCTCATGA   107

       *       *   **    * *    **  *     ** **      * **
bTHP  GAAC  CCAT   GAG GATGG  AACTTGCTTTGGAACTGAGTGGCCACAAGTATACA   167
rTHP  AATC  TGGT   GAA GCT G  TTCTT CAGATCAGGTGTGTGTTC CAGGCTTCGA   156
hTHP GAATGGGTGTTTCAGAAGGGT GCCCTCTC CAAGACAGGTGCACCTCC CATCTGGGGC   164

         **      *  *   *      *       **  ** *   ** * *     ** *
bTHP TCCCAGTGGGGACAGTGAGC  ACCCCTTTTCTCCTGGAGCAGCCTGGCTTCAGATTCTG   225
rTHP AGCAAAT GTTTCTGTTATCCTAACCCAGGCTGGCTTCAG ATATTGTCTT TTTTCCTG   213
hTHP AGTGAAT  ATCCTTTTGTCCTTATGCAGCCTGGCTTCAG ATACTGGC T TCTGCCTG   219

        *  **   *  ****  **  ***  *     ***** * ********   *****
bTHP  GCCTCTG  CTTTCCACTTTGTGCTTTTCAATGACCAAG AAAAGCCCAGGCACTTGGA   281
rTHP CCCCTTTGGTATTTCCACCTTGTCCTTGCCCAGGTCCAAGAAAAAGCCCAGAACCTTGGC   273
hTHP GCTCCTTG  AT CCCACCCTGCCCTTGTCAGTGACCAAGAAGAAGCCCAGCACCTTGGC   276

     * *  **  ********* ******    **   *** * **  * *** *    *****
bTHP ATTTTTACCCAGTTAATTTCTAACTAAAGAACCTCTCGTTGCCAAAAGATATA AAACA   340
rTHP ACTGCTTTGCCAGTTAATGTCTAACCGAGGAATGTCTTGCTGCCAAAAGGT GC AAACA   331
hTHP ACTGCTTTCCCAGTTAATTTCTAACTATGGAATCTCTTGCTGTTAGAAGGT GCGAAACA   335

     * * *******  **  ******  *****  ***** ************    ******
bTHP GAGCCCTTGTAACTCTGGGCACAACTGTGA  CCCCAGTGTCAATCATTTGGGGTCTCTA   398
rTHP GAGACCTTGTATTTCCAGGCACAGGTGTGA CCCCCAATGTCAATCATTTTCTGTCTCTA   390
hTHP GTGACCTTGTATTTCCGGGCACAGGTGTGACCCCCCAATGTCAATCATTTGGGGTCTCTA   395

      ***     ******   *********  *****     * ** ****  ***  * **
bTHP CCTATTAG GGAAAA  GAACAACAACCACCTCACAGCCTAGAAAAGGAAAACACTGTGT   455
rTHP ACTACCAGAGGAAAAACTAACAACAACAGCCTCATGGTTTGGAAAAGGTGAACTCTATGC   450
hTHP GCTA  TTAGGAAAAA GAACAACAACAACCTCACAGCTTGGACAAGGCAAACATTATGC   452

     * *   ****  **** *** *******   *** **  * ** *  *****  ****
bTHP CAAAAGGAA AAATATTC CACCCCCATTAAAATAA  TTAAGAAACAGAACCAGAGGA   511
rTHP CAAATGGGAAGAAAAATTCTGACCCCCACAGAAACAATCTCAAGAGGCAGAAGCAGAGAA   510
hTHP C AGGAGGAAAAAATATTC CACCCCCAAGAAAACAATATCAAAAAACAGAACTAGAGAC   510

     * ****** * ** *  * * ***   **   *    * ******  **        **
bTHP TCATTGGA G GAAA  GACTGCCAGTGGGGGACAGATGTATATATATAGATATGATAGT   567
rTHP TAATTGGAGG GAGAGGGCCAGCCAAGGGCAGACATATATATATATATATTGATCACAGG   569
hTHP TAATTGGAGGAGAGATTGCCAGCCTGGGGCAAATGTGTATATATAAGTA TG    AGG   564

     *** *  *    *    *    *         **** ********** ** ****  **
bTHP CACCTACTTGTAAAAGGATTAATTCTACCTTTCTGGTTTCAGGTAAGGCTATCTGCAGA   626
rTHP CACTTACTTGTGAATGGACCAGTCCT   GTCCTGGGTTCAGGTAAGACTGTCTGGAGC   625
hTHP CACATCATCACCAGACTAACTCTACC   TTTCTGGCTTCAGGTAAGGCTATCTGTAGC   620
```

FACING PAGE
FIG. 5. Sequence alignment of promoters from three species: rat, cow, and human. The start sites are circled. TATA and inverted CCAAT boxes have been underlined, as have conserved regions at nucleotides 258–300, 359–400, and 468–487.

outer medulla to enrich for tissue-containing THP message at high amounts, and we have recently also successfully produced in vitro transcription competent extracts from this source (data not shown). The kidney nuclei served a twofold purpose. First, the nuclei are a source of the THP gene in its natural chromatin context; this allowed us to identify a kidney-specific DNase I hypersensitive site and to thereby locate the promoter. Second, the nuclei are a source of nuclear protein that could be used to carry out DNA binding studies (data not shown) and in vitro transcription.

The identification of the cap site proved complicated. The rat cDNA clone rTHP-1 ends to 16 bp upstream of the initiator methionine codon. Primer extension with two gene-specific primers indicated that the rat cap site was 68 bp upstream of the initiator codon. The rat extension product was not likely to be contiguous with the genomic sequence due to two reasons: (1) RNase protection analysis using an antisense riboprobe in this region gave a protected fragment that had its 5′ end only 39 bp upstream of the initiator codon (data not shown); this region has a consensus splice acceptor sequence; (2) the sequences upstream of the putative cap site—were the primer extension product to be contiguous with the genomic sequence—might have contained easily identifiable promoter elements. Because this was not the case, it seemed likely that a small exon having a maximum length of 30 bp existed somewhere upstream in the gene.

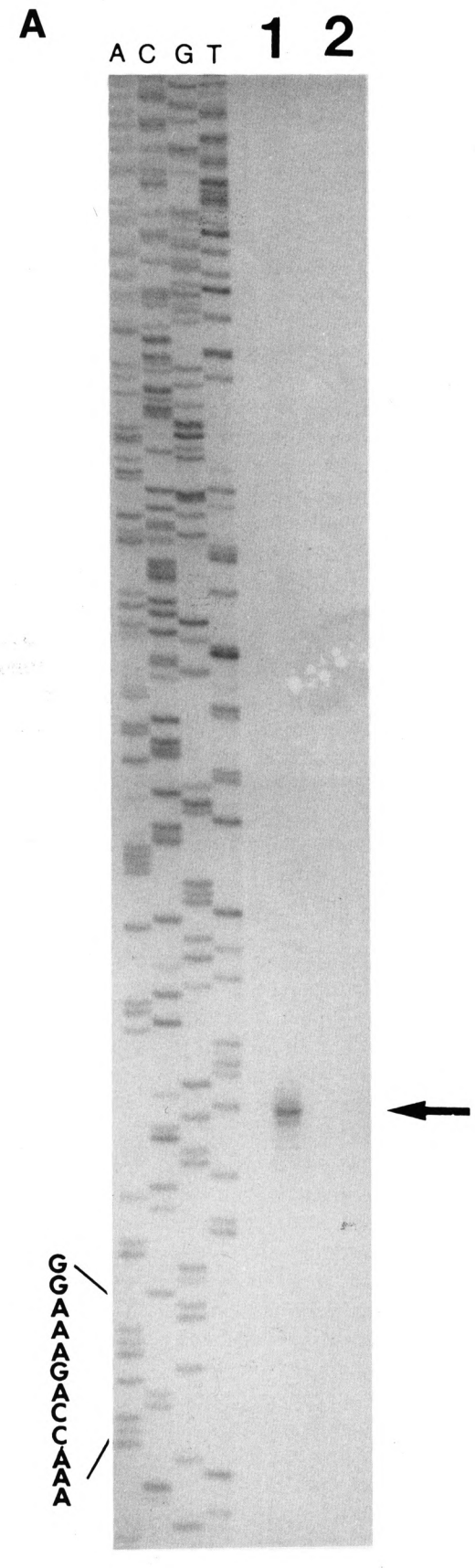
Pennica et al. (1987) isolated two classes of human THP cDNAs. The first class has a 25 bp exon that they refer to as exon "a" at the 5′ end. The second class has as its 5′ end the region just upstream of and contiguous with the second exon of the first class of cDNA. They refer to this as exon "b" and propose an alternative splicing event that gives rise to the two classes. We performed primer extension analysis using a human exon 2-specific primer on human RNA and found two extension products that agree with the boundaries of the cDNA ends of Pennica et al. (data not shown). However, both the rat and bovine THP mRNAs are associated with only one 5′ end. We think that this species-specific variation arises because a fraction of the human THP RNAs retain a region as the 5′ end (due to an abnormal splicing event) that in the other two species is an intron. We favor this explanation over the possibility that the human gene is associated with an additional promoter located upstream of exon "b" because there are no well-conserved promoter elements there and because that region does not serve as a promoter in vitro (data not shown).

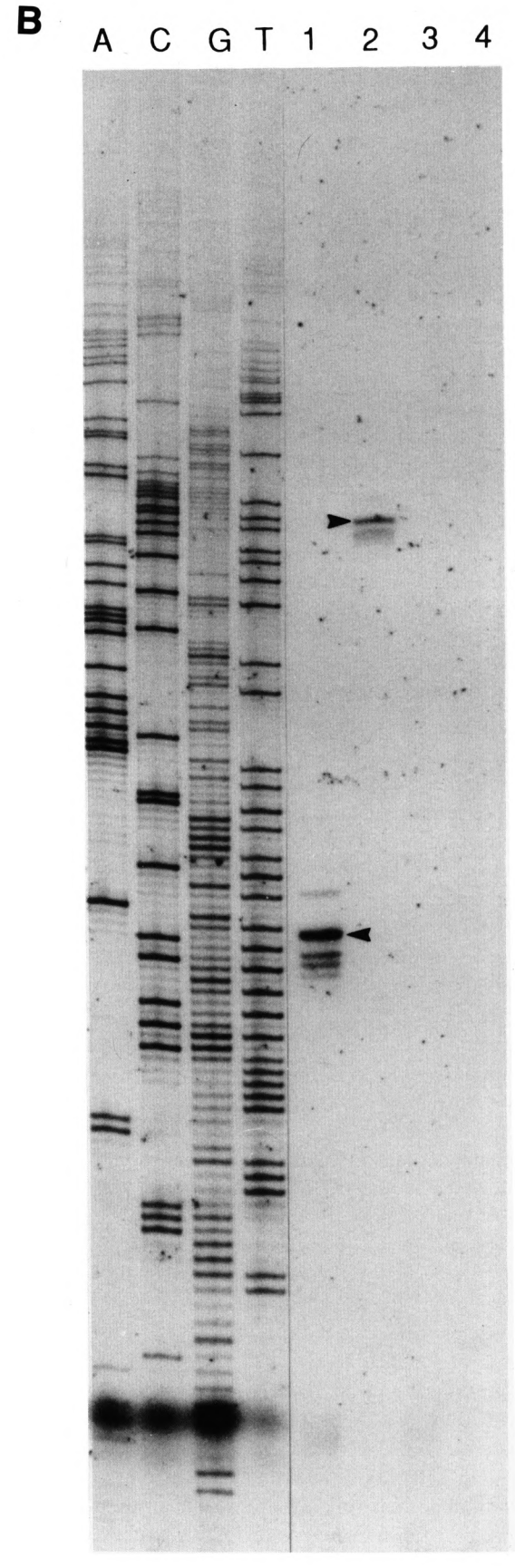
For these reasons, it seemed likely that the expected 30 bp first exon in the rat gene would be the homologue of the human exon "a." If this were indeed the case, we expected extensive interspecies homology in the putative promoter regions upstream of these first exons. However, it proved difficult to locate the homologue of exon "a" on the rat gene by hybridization, due to the small size of the oligomer and its strong hairpin structure. Instead, the clue to the location of the rat homologue of exon "a" came from the mapping of a kidney-specific DNase I hypersensitive site, located approximately 700 bp upstream of the splice acceptor site of the rat exon 2. Importantly, DNA sequencing of a 350 bp region covering the hypersensitive site revealed an 80% homology with the genomic sequence upstream of human exon "a." The bovine gene contains the same conserved 350 bp region, and the homology between the bovine and the rat region is 72% (see Fig. 5).

RNase protection analysis identified a 5′ end in the rat gene that agreed precisely with the primer extension result and that accounted for the 30 additional bases (see Fig. 7). The bovine primer extension and RNase protection results further agreed with the rat RNA mapping.

The promoter region is associated with a TATA element at −34 and an inverted CCAAT element at −65. In addition, three highly conserved elements were seen at rodent nucleotides 258–300, 359–400, and 468–487 (underlined in Fig. 5).

We wished to see whether the THP promoter would function in an in vitro transcription system because the absence of a cell line prevented us from transient assay testing. A G-less cassette reporter (pTHP 380) gave an accurate size transcript of 380 bp. Furthermore, the inclusion of α-amanitin at 2.5 μg/ml completely abolished transcription, as would be expected of a Pol II promoter. These results show that the conserved region can serve faithfully as a Pol II promoter in an in vitro system. Future experiments will examine the role of the individual conserved elements through the use of this in vitro system. Ultimately, reporter constructs in transgenic mice will reveal the precise elements required for high-level, kidney-specific expression of the THP


A
A C G T
1 2
GGAAAGACCAAA


B
A C G T 1 2 3 4

FACING PAGE

FIG. 6. (A) Primer extension of bovine kidney outer medullar RNA (20 μg) (lane 1) and tRNA (lane 2) with primers, bTHP4 and bTHP7 (not shown), randomly chosen from exon 2. The major extension product identified is 16 nucleotides upstream from the 5′ end of bTHP1. A sequencing ladder was produced with bTHP1 and the same oligo primers. (B) Primer extension analysis of rat kidney mRNA (2 μg) with primers RTHPRE36 (lane 1) and RTH5 (lane 2). Lanes 3 and 4 are extensions using rat liver mRNA (2 μg) and primers RTHPRE36 and RTH5, respectively. The major extension products (102 nt in lane 1 and 159 nt in lane 2) are marked by arrows. A sequencing ladder was run next to the reactions for sizing purposes.

message. With the cloning of THP genomic fragments containing all the exons as well as the promoter and considerable 5′ flanking sequence, we anticipate performing these experiments in the future. Finally, because virtually nothing is known about kidney-specific gene expression and nephron segment-specific expression, the studies reported here initiate a data base on this topic with implications to understanding kidney growth and differentiation.

ACKNOWLEDGEMENT

This work was supported by NIH DK44921 to V.P.S.

REFERENCES

R. Breatnach and P. Chambon (1981), Annu Rev Biochem 50, 349.

P. Chomczynski and N. Sacchi (1987), Anal Biochem 162, 156–159.

Current Protocols in Molecular Biology (F. Ausubel, ed.) (1991), Green Publishing Associates and Wiley-Interscience.

V. DeSimone and R. Cortese (1988), Oxford Surveys on Eukaryotic Genes 5, 51–90.

A. P. Feinberg and B. Vogelstein (1983), Anal Biochem 132, 6–13.

A. P. Fletcher (1972), in Glycoproteins: Their Composition, Structure and Function (A. Gottschalk, ed.), Elsevier, New York, pp. 892–908.

K. Gorski, M. Carneiro, and U. Schliber (1986), Cell 47, 767–776.

D. S. Gross and W. T. Garrard (1988), Annu Rev Biochem 57, 159–197.

C. Hession, J. M. Decker, A. P. Sherblom, S. Kumar, C. C. Yue, R. J. Mattaliau, R. Tizard, E. Kawastima, S. Scheissmer, S. Heletky, E. P. Chow, C. A. Burne, A. Shaw, and A. V. Muchmore (1987), Science 237, 1479–1484.

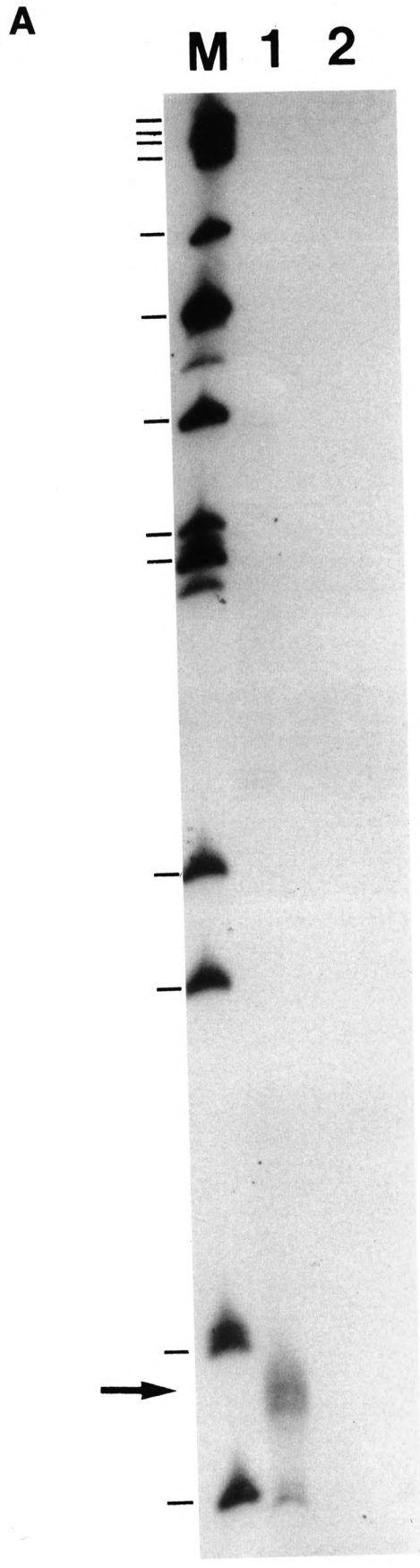
A. V. Muchmore and J. M. Decker (1985), Science 229, 479.

D. Pennica, W. J. Kohr, W. Kuang, D. Glaister, B. B. Aggarwal, E. Y. Chen, and D. V. Goedder (1987), Science 236, 83–88.

J. Sambrook, E. F. Fritsch, T. Maniatis (1989), Molecular Cloning, A Laboratory Manual, 2nd ed. (N. Ford, ed.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

E. A. Schenk, P. H. Schwartz, R. A. Lewis (1971), Lab Invest 25, 92.

I. Tamm and F. L. Horsfall (1950), Proc Soc Exp Biol Med 74, 108.



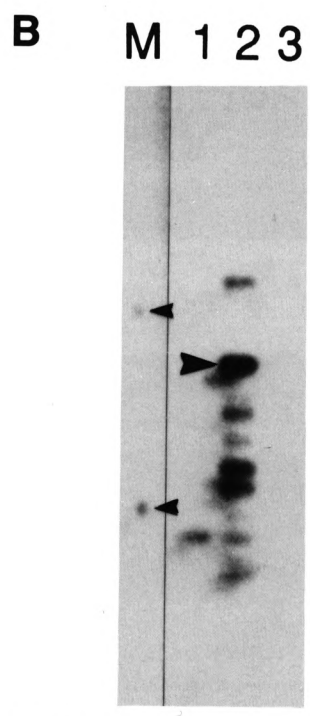


FIG. 7. (A) RNase protection assay for the bovine THP. Lane 1: bovine outer medulla RNA. Lane 2: tRNA. EcoRI-Pst I 800 bp genomic fragment that contains the first exon was subcloned into pBluescript II. The plasmid was digested with Ssp I, and the 220 bp antisense fragment was transcribed and used for hybridization. The arrow indicates a DNA nucleotide size range of 32–33, consistent with the first exon size of 31 nt, as shown by the primer extension assay. Bars indicate nucleotide sizes of 711, 489, 404, 364, 242, 190, 147, 118, 110, 67, 57, 34, 26, from the top. (B) Rnase protection analysis of rat samples using an antisense riboprobe from −350 to +35. Lane 1: liver total RNA (10 μg). Lane 2: kidney total RNA (10 μg). Lane 3: 10 μg tRNA. The top arrow in the marker lane points to 34 nt; the bottom arrow to 26 nt. The major transcript end (arrow in lane 2) is at 30 nt.

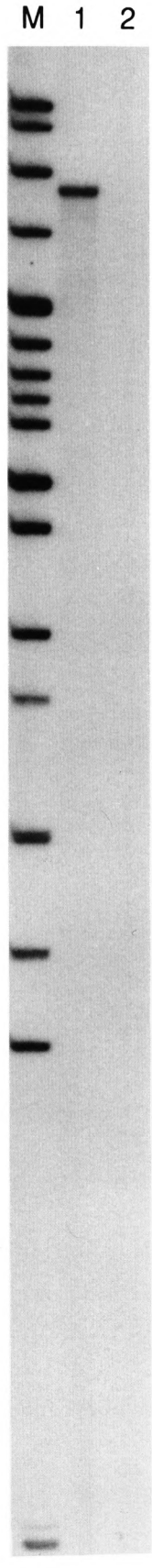


FIG. 8. In vitro transcription using kidney nuclear extracts. A total of 1 μg of template and 60 μg of nuclear protein were used for both reactions. Lane 1: pTHP380. Lane 2: pTHP380 with 2.5 μg/ml α-amanitin.